

**Modelling the effect of in-stream and overland dispersal on gene flow in river networks**

A. Chaput-Bardy<sup>1†</sup>, C. Fleurant<sup>2\*</sup>, C. Lemaire<sup>1‡</sup> and J. Secondi<sup>1</sup>

<sup>1</sup> Laboratoire d'Etudes Environnementales des Systèmes Anthropisés, Université d'Angers, 2 bd Lavoisier, 49045 Angers cedex 01.

<sup>2</sup> Unité de Recherche Paysage, AgroCampus Ouest, 2 rue Le Nôtre, 49045 Angers cedex 01.

† Current adress: UMR 7179 CNRS-MNHN, Equipe EVOLTRAIT, 4 av du Petit Château, 91800 Brunoy.

‡ Current address: Laboratoire 2EPF, UMR PaVé-INRA-Agrocampus Ouest-Université d'Angers, 42, rue Georges Morel, BP 57 49071 Beaucouzé, Cedex.

Keywords: genetic structure, individual-based model, freshwater organisms, simulations

\*Corresponding author: Cyril Fleurant

Unité de Recherche Paysage, AgroCampus Ouest, 2 rue Le Nôtre, 49045 Angers cedex 01, France

Fax: +33 2 41 22 55 28

E-mail address: [cyril.fleurant@agrocampus-ouest.fr](mailto:cyril.fleurant@agrocampus-ouest.fr)

## ABSTRACT

Modelling gene flow across natural landscapes is a current challenge of population genetics. Models are essential to make clear predictions about conditions that cause genetic differentiation or maintain connectivity between populations. River networks are a special case of landscape matrix. They represent stretches of habitat connected according to a branching pattern where dispersal is usually limited to upstream or downstream movements. Because of their peculiar topology, and the increasing concern about conservation issues in hydrosystems, there has been a recent revival of interest in modelling dispersal in river networks. Network complexity has been shown to influence global population differentiation. However, geometric characteristics are likely to interact with the way individuals move across space. Studies have focused on in-stream movements. None of the work published so far took into consideration the ability of many species to disperse overland between branches of the same network though. We predicted that the relative contribution of these two dispersal modalities (in-stream and overland) would affect the overall genetic structure. We simulated dispersal in synthetic river networks using an individual-based model. We tested the effect of dispersal modalities, i.e. the ratio of overland / in-stream dispersal, and two geometric parameters, bifurcation angle between branches and network complexity. Data revealed that if geometrical parameters affected population differentiation, dispersal parameters had the strongest effect. Interestingly, we observed a quadratic relationship between  $p$  the proportion of overland dispersers and population differentiation. We interpret this U-shape pattern as a balance between isolation by distance caused by in-stream movements at low values of  $p$  and intense migrant exchanges within the same branching unit at high values of  $p$ . Our study is the first attempt to model out-of-network movements. It clearly shows that both geometric and dispersal parameters interact. Both should be taken into consideration in order to refine predictions about dispersal and gene flow in river network.

## 1. Introduction

Dispersal is a key life history trait for population processes as it contributes to gene flow, prevents local extinction, and determines the probability of patch recolonization (Hanski, 1999; Clobert et al., 2001; Bullock et al., 2002; Bowler and Benton, 2005; Ronce, 2007). The spatial distribution of dispersal events across an area determines preferential connections and gene flow level between patches. Connectivity level between demes affects migration-drift equilibrium and the overall level of genetic differentiation of populations. Modelling dispersal in landscapes is thus an essential step to investigate the relationships between the distribution of habitat patches, connectivity networks and the resulting spatial genetic structure of populations (Campbell et al. 2007). Model outputs help to build predictions for instance about the effect of connectivity on the inbreeding level and the potential for local adaptation of demes which are both major issues from a theoretical and conservation point of view.

Natural environments are anisotropic which means that heterogeneity in landscape composition and spatial organisation causes individuals to move along preferential pathways across the landscape matrix (Gustafson and Gardner, 1996; Béliste, 2005). Many studies attempted to model gene flow in different types of landscapes (Nuismer et al., 2003; Cushman et al., 2006; Shaw et al., 2006; Ceddia et al., 2007; Vuilleumier and Fontanillas, 2007). Still, river networks can be considered as a particular case of landscape matrix (Campbell et al. 2007). The peculiarity arises from their topology. River networks are continuous stretches of habitats, suitable for breeding or dispersal, connected in a dendritic (i.e. branching) pattern. The “biological” distance between two sites is not defined by the straight-line distance (Euclidian distance) but by the length of the watercourse between the two points (Fagan, 2002). Consequently, the geometrical configuration of the river network determines the distribution of distance between breeding sites. The distance following the watercourse (Little et al., 1997) and the number of branches (Gardner et al., 2003) are parameters that could affect dispersal of individuals, gene flow patterns and ultimately spatial genetic structure. However, in river networks, individuals do not necessarily disperse along a watercourse (i.e. in-stream dispersal). They can also move between branches (i.e. overland dispersal) if all life stages are nor strictly aquatic, or if floods connect different watercourses.

In species whose dispersal is restricted to the watercourse, movements are constrained by water flow, network topology, and physical barriers to dispersal such as bordering terrestrial habitat (Slatkin, 1985; Rissler et al., 2004; Lowe et al., 2006). It results from these

properties that gene flow occurs preferentially along branches (Ward et al., 1994; Bilton et al., 2001; Beck and Pruett-Jones, 2002; Power and Dietrich, 2002; Peckarsky et al., 2005). Gene flow is generally asymmetric as more individuals move downstream following the water flow than the other way round (Hernandez-Martich et al., 1995). In addition, the hierarchical structure of river networks tends to isolate populations living in different branches from each others (McGlashan et al., 2001). Individuals are then distributed among different watercourses when they move upstream, whereas they are concentrated, joining at confluences, when they move downstream (Power and Dietrich, 2002). These two types of movement are expected to result in decreased genetic diversity upstream (Schmidt et al., 1995) and increased genetic differentiation between branches (Power and Dietrich, 2002). In addition, upstream populations are more sensitive to genetic drift because of their geographic isolation (Fraser et al., 2004). Therefore, the genetic structure of populations is supposed to reflect the hierarchical structure of the river network (Meffe and Vrijenhoek, 1988). This assumption was verified in 56% of fish species (A vise, 2000), the freshwater prawn, *Macrobrachium australiense* (Cook et al., 2002), and the white-clawed crayfish *Austropotamobius pallipes* (Grandjean et al., 1997).

Nevertheless, genetic patterns do not always match the spatial organization of the river network because of barriers to dispersal caused by dams (Hurwood and Hughes, 1998; McGlashan and Hughes, 2000) or out-of-network dispersal (Hurwood and Hughes, 2001). A number of species like freshwater insects (Kovats et al., 1996; Miller et al., 2002; Petersen et al., 2004; Macneale et al., 2005) or amphibians (Lowe, 2002) shows overland and in-stream dispersal. For these organisms, overall genetic differentiation is supposed to decrease with the level of overland connectivity between nearby but independent watercourses (Gibbs et al., 1998). This prediction has recently received support from an empirical study in the Banded damselfly, *Calopteryx splendens* (Chaput-Bardy et al., 2008). Genetic analyses revealed that gene flow preferentially occurred along watercourses. However, overland dispersal accounted for the mismatch between the spatial genetic structure and the river network structure.

Two network parameters are likely to influence the effect of overland dispersal on genetic structure. (i) The mean angle between two branches joining at a confluence is also of importance as the Euclidean distance between two sites of adjacent branches decreases with the bifurcation angle. (ii) For a given area, the number of branches that receive migrants from other branches increases with the complexity of the network, i.e. the number of branches, owing to the higher density of watercourses. For organisms with strict in-stream dispersal,

higher complexity is expected to promote differentiation between regions of the network (sub-catchments). In contrast, it is expected to lower population differentiation when overland dispersal allows gene flow between otherwise distant branches. In genetic terms, overall population differentiation is expected to be positively correlated with the mean bifurcation angle but the relationship with network complexity is less straightforward.

In summary, it can be predicted that the population genetic structure in a river network depends on the geometry of the river network, the dispersal modalities used by the species, and the presence of barriers. The first mathematical models investigating the effect of river network topology on genetic patterns were carried out a few decades ago (Sawyer, 1978; Meffe and Vrijenhoek, 1988). However, a revival of interest in the effect of river structure on dispersal or gene flow has been occurring (Charles et al., 2000; McGlashan and Hughes, 2000; Fagan, 2002; Lowe, 2002; Neuenschwanger, 2006; Labonne et al., 2008; Cote et al., 2009). Recently, Labonne et al. (2008) demonstrated using simulations the effect of the size and the number of branches on connectivity as well as on metapopulation persistence. However, none of these authors attempted to model out-of-network dispersal. The effects of bifurcation angles and network complexity on genetic structure remain to be investigated for species or life stages able of out-of-network dispersal (Campbell Grant et al., 2007) (Fig. 1). In the present study, we used an individual based model to simulate in-stream and overland dispersal and tested the effects of the branching pattern, bifurcation angles and the probability of overland dispersal on the overall genetic structure of demes in the network.

## 2. Materials and methods

### 2. 1. Modelling of synthetic river networks

River networks have two main properties. Firstly, they are made of segments (watercourse stretches) and confluence points that form a branching network (Ganio et al., 2005). The geometry of a network is characterized by the hierarchical structure of watercourses (segments) categorized in orders. Many classification systems have been put forward but we decided upon Strahler's system (Strahler, 1957) that is the most widely used. The classification system is as follows: (i) headwaters are considered as first order stream segments, (ii) when two stream segments within the same order  $\omega = i$  merge, the stream segment resulting from this confluence is considered as order  $\omega = i+1$ , (iii) when two stream segments of different orders,  $\omega = i$  and  $\omega = j$  merge, the stream segment resulting from this confluence is of order  $\omega = \max(i,j)$  (Strahler, 1957). The river network order corresponds to the highest index value of a segment. This classification puts forward general geometric laws. Among them, Horton's laws (Horton, 1945) describe the way stream networks are organized. These laws express the so-called bifurcation ratio  $R_B$  and length ratio  $R_L$ , also known as Horton's ratios. A great number of experimental studies on stream networks (Tarboton et al., 1990; Rosso et al., 1991) revealed that these ratios are rather stable and fluctuate between 3 and 5 for  $R_B$  and between 1.5 and 3.5 for  $R_L$ . Horton's laws also make it possible to work out the  $R_B$  and  $R_L$  (equation 2), where  $\bar{l}_k$  is the average value of the morphometric lengths of  $k$  order and  $N_k$  is the number of morphometric lengths of  $k$  order.

$$R_L = \frac{\bar{l}_k}{\bar{l}_{k-1}}, \quad R_B = \frac{N_{k-1}}{N_k} \quad (2)$$

### 2. 2. Modelling in-stream and overland dispersal in synthetic river networks

In our synthetic networks, each node is connected to one or two upper nodes (upstream) and one lower node (downstream). Upstream and downstream dispersal is allowed. Dispersal is defined as the fraction  $m$  of adults that move out of their natal node before breeding, so it is different from effective migration. Among dispersers, a fraction  $(1-p)$  of adults disperse in-stream, i.e. along the watercourse, and a fraction  $(p)$  disperses overland, i.e. between branches

of the network. River networks are discretized into evenly spaced nodes. Each node is a potential breeding site. Individuals can disperse up to four nodes when moving in-stream, and to an equivalent distance when moving overland to another branch. The dispersal distribution is uniform so that individuals can reach any node within a 4-node range. We allowed  $p$  to vary between 0 and 1. A zero value indicates that only in-stream dispersal occurs. When  $p=1$ , all adults disperse overland when possible. As a general rule for  $p>0$ , individuals disperse in-stream if no site can be reached by an out-of network movement. Bifurcation angle  $\alpha$  determines the overland distance between two nodes on adjacent branches but not their in-stream distance. Successful overland dispersal is more likely for low values of  $\alpha$  as distances between branches are shorter. On the contrary, in-stream dispersal should prevail in network with high values of  $\alpha$ .

Boundary nodes (boundaries of the river catchment) are linked to upper nil nodes. For these nodes we consider that individuals disperse out of the network. The reason is that boundary nodes receive fewer migrants (only from one direction) and are thus more prone to genetic drift. They could thus bias the estimation of the overall population structure.

### *2. 3. Colonization, demographic and genetic parameters*

Individuals were introduced from the lowermost node. We allowed colonization of the entire network from this single node. The carrying capacity  $K$  for each node was fixed to 100 adults. If the number of adults after migration in a population exceeded  $K$ , then a sample of  $K$  individuals was randomly drawn from the pool of resident and migrant adults. Extra individuals were not allowed to breed. Strictly monogamous pairs were formed randomly from the sample of adults. Each pair produced 5 offspring so that no variation in breeding success occurred. There was no overlap between generations.

Individuals were characterized by ten bi-allelic loci with a mutation rate of  $10^{-8}$  like SNP (Single Nucleotide Polymorphism) markers (Brumfield et al., 2003). A population, defined as all adults present on a given node. Sex of new individuals were drawn at random. Individuals of the initial sample (first generation on the lower node) were heterozygous for the ten loci. By this way all simulations start with the same initial conditions : allele frequencies set at 0.5 for the ten loci. For all subsequent generations and populations egg genotype resulted from the random sampling of one allele from each parent. Hardy-Weinberg equilibrium was assumed for each population.

#### 2. 4. Test of network geometry and dispersal characteristics on genetic structure

We modeled three river networks representative of many catchments observed in natural environment. The Horton parameters for the selected networks were as follows: Network A  $\omega=3, R_L=2, R_B=2$ ; Network B  $\omega=3, R_L=2, R_B=3$ ; and Network C  $\omega=3, R_L=2, R_B=4$ . For each of these three river networks, we created six additional networks with different values of  $\alpha$ , the bifurcation angle between two branches joining at a confluence. We set  $\alpha$  respectively to 15, 30, 45, 60, 75, and 90 degrees. Higher values are not expected to occur frequent. Keeping  $\alpha$  constant across the whole network could cause crossings between higher order branches. In order to avoid this problem, we used a simple rule. Each bifurcation angle involving one last order segment was set to  $\alpha$ . Angles for all other bifurcations were set to  $\alpha/2$ .

A total of 18 networks were synthesized, the Fig. 2 represents some of them. Size was fixed so that the distance between the lower and upper node was 26 nodes. Network size, i.e. the number of nodes, was 54 for  $R_B=2$ , 88 for  $R_B=3$ , and 132 for  $R_B=4$ .

For each network we selected six values of  $p$ , the fraction of overland dispersers: 0, 0.2, 0.4, 0.6, 0.8, and 1. For all 36 combinations of  $\alpha$  and  $p$  we ran 20 replicates. In addition, we replicated these simulations for all parameter combinations at two dispersal levels ( $m=0.1$  and 0.2). Each simulation was run for 300 generations. Preliminary tests using the more complex network (Network C) showed that this length was sufficient to reach migration-drift equilibrium for the range of network sizes we simulated (Fig. 3). Overall, we ran 4320 simulations ( $3 \text{ networks} \times 6\alpha \times 6p \times 2m \times 20 \text{ replicates}$ ).

We computed the mean and the variance of the overall  $F_{ST}$  value for each for  $\alpha$  and  $p$  combinations and  $m$  values. We thus included all populations and all individuals in the computation of the  $F$  statistics. For each network, we visualized the effect of  $\alpha$  and  $p$  on  $F_{ST}$  and  $F_{ST}$  standard deviation as a response surface fitted using the weighted least square method using STATISTICA 7.1 (StatSoft, Inc, 1984-2005).

Generalized linear models (GLMs) were constructed to test for the influence of  $R_B$ ,  $\alpha$ ,  $p$  and their interactions on  $F_{ST}$ . We carried out a factorial analysis including all factors and their pairwise interactions. We used a squared term of  $p$  to test for non-linear effect of this variable (McCullagh and Nelder, 1983). GLMs assumed a normal error structure, as the response variable ( $\log F_{ST}$ ) was adjusted to a Gaussian distribution. A stepwise Akaike's Information Criterion with a backwards selection procedure was used to select the most



parsimonious model from the complete model:  $\log(F_{ST}) \sim R_B + \alpha + m + p + p^2 + R_B:\alpha + R_B:m + R_B:p + R_B:p^2 + \alpha:m + \alpha:p + \alpha:p^2 + m:p + m:p^2 + p:p$  (Burnham and Anderson, 2002). For each model considered we also calculated the percentage deviance explained (%DE) as a measure of goodness-of-fit. Model selection was carried out using the StepAIC function in MASS package available in R (R project for statistical computing, ver. 2.8.1, <http://www.r-project.org/>).

### 3. Results

#### *Connectivity and dispersal modality*

Figure 4 shows that, in synthetic river networks, the distribution of pairwise distances differs for the two modalities of dispersal considered. The probability of reaching a node is higher for out-of-the-network dispersal than for in-stream dispersal for all distances (Fig. 4). In addition, the probability difference between in-stream and overland dispersal increased with  $R_B$ . Finally, the difference depended on the bifurcation angle as it increased with  $\alpha$  for a given  $R_B$  (Fig. 4). Thus, this graph shows that the level of network complexity  $R_B$  influences the probability of successful dispersal. It is also indicative of an interaction between network geometry,  $R_B$  and  $\alpha$ , and dispersal modalities on the biological connectivity of the whole river network.

#### *Effect of dispersal parameters $m$ and $p$*

The complete model,  $\log(F_{ST}) \sim R_B + \alpha + m + p + p^2 + R_B:\alpha + R_B:m + R_B:p + R_B:p^2 + \alpha:m + \alpha:p + \alpha:p^2 + m:p + m:p^2 + p:p^2$ , gave a good fit to the dataset (residual adjustment test,  $p > 0.1$ ). According to the model selection procedure (Stepwise AIC), the two best models accounted for more than 86.5% of explained deviance (Table 1). The best model, Model 1, contained all variables ( $R_B$ ,  $\alpha$ ,  $m$ ,  $p$  and  $p^2$ ) and all but one interaction  $\alpha:m$ , and the second best model, Model 2, was the complete model (Table 1). Results of both models were strongly similar; consequently we chose to show the results for Model 1 only. Most interactions were significant although their residual deviance was small (Table 2). The large sample size of simulations might have caused the undue detection of small effects without little biological sense. We thus focused our analysis on the main significant effects and interactions.

The main factors influencing  $F_{ST}$  were the migration rate  $m$  (residual deviance = 83.49,  $p < 0.0001$ , table 2), and squared overland dispersal  $p^2$  (residual deviance = 31.78,  $p < 0.0001$ ) (Table 2, Fig. 6). Dispersal rate  $m$  had the most influential effect on  $F_{ST}$  (Table 3). For  $m=0.20$  the response surface was strongly flattened, which outlined the fact that  $F_{ST}$  weakly responded to variations of the bifurcation angle  $\alpha$ , and the proportion of overland dispersers  $p$  when  $m$  was high enough (Fig. 6). Surprisingly, we found a quadratic relationship between  $F_{ST}$  and  $p$ .  $F_{ST}$  decreased for values of  $p$  between 0 and 0.5. It increased for values of  $p$  above 0.5 (Fig. 6). A similar relationship was logically observed for the standard deviation of  $F_{ST}$  (Fig. 7). Thus, a minimum of genetic differentiation was reached for intermediate values of  $p$ .

In order to make sure that the quadratic relationship was not an artefact caused by dispersal rules, we ran additional simulations where overland disperser died when they could not reach a node. We carried out ten replicates for two values of the bifurcation angle ( $\alpha=15$  and  $90^\circ$ ) on the following network:  $\omega=3$ ,  $R_B=3$ ,  $m=0.1$ . Results showed again a quadratic relationship between  $p$  and  $F_{ST}$  (data not shown). The only difference was null  $F_{ST}$  values for  $p=1$  as individuals could not disperse overland out of the initial node and thus could not colonize the network.

#### *Effect of network parameters*

The geometry of the network (i.e. the number of branches and the bifurcation angle between branches) significantly affected the overall genetic structure. According to the positive values of their coefficients, an increase in  $R_B$  or  $\alpha$  was reflected by an increase in  $F_{ST}$  (Table 3). These two factors respectively took 3.99 ( $p < 0.0001$ ) and 7.38 ( $p < 0.0001$ ) of the residual deviance (Table 2). Fig. 7 illustrates the overall effect of  $\alpha$  on  $F_{ST}$ . When overland dispersal occurs, overall population differentiation increases with distance between branches. However, Fig. 6 and Table 2 show that the level of network complexity  $R_B$  influenced the relationship between the two variables ( $R_B:\alpha$ , residual deviance = 1.05,  $p < 0.0001$ ). Similarly,  $F_{ST}$  increased with bifurcation ratio. Thus, population genetic differentiation increased with network complexity  $R_B$ . Here again, the relationship between this variable and  $F_{ST}$  was not strictly linear but strongly affected by the interaction with  $p$  the proportion of overland dispersers (residual deviance = 5.61,  $p < 0.0001$ ).

## 4. Discussion

We investigated the effects of network geometry and dispersal modalities on the genetic differentiation of populations. To our knowledge, this study is the first attempt to explicitly consider out-of network dispersal, i.e. movements of individuals between branches. Results clearly showed that dispersal and network characteristics not only influence the overall level of population differentiation but the way they interact is important too. As a consequence, both factor types should be considered jointly when investigating population processes in river networks.

### *Effect of network parameters*

Variation of the bifurcation ratio  $R_B$  strongly influences the density of watercourses when keeping constant the catchment area. An individual dispersing out-of-the network has thus a better chance to reach another branch in highly than in poorly ramified networks. Fig. 3 and 4 show that the number of nodes that can be reached is higher when dispersing overland than when dispersing in-stream, and that the difference tends to increase with  $R_B$ . They clearly illustrate the importance of considering the different dispersal pathways available to individuals.

Simulations revealed that overall  $F_{ST}$  was positively related to  $R_B$ . For a given dispersal range, the fraction of movements within branches of the same branching unit increased with the number of branches. In biological terms, this means that differentiation between sub-catchments is favored when network complexity increases. As predicted the bifurcation angle  $\alpha$  affected the genetic structure too. Individuals have better opportunities to successfully disperse overland in network with low values of  $\alpha$  because of the reduced mean pairwise distance between sites. Consistently, we observed a positive relationship between  $\alpha$  and the overall level of population differentiation. Thus, the simulation results support our initial predictions about the relationships between river network parameters and the level of population differentiation.  $F_{ST}$  tended to increase with network complexity and bifurcation angle between branches, i.e. with decreasing opportunity for successful overland dispersal. However, the effect of network parameters did not only consist in linear changes of the overall  $F_{ST}$  value. Interactions between  $R_B$  and  $p$  contributed to the bending of the surface response.

## *Effect of dispersal parameters $m$ and $p$*

Dispersal  $m$  and the proportion of overland dispersers  $p$  appeared as the major factors contributing to population differentiation. The reduction in deviance when these factors were considered was much greater than for any other factor. Dispersal alone seemed to be the most influential on  $F_{ST}$ , as shown on Fig. 4. Doubling its value from 0.1 to 0.2 caused a strong flattening of the surface response whatever the combinations of the other parameters. Such a result suggests that moderate to strong migration would cancel the effect of network geometry. Thus, the effects of other factors would only be detected under a range of low dispersal rates.

Unexpectedly, we observed a quadratic relationship between  $p$ , the proportion of overland dispersers, and population differentiation  $F_{ST}$ . The effect was reinforced with increasing values of  $R_B$ . We interpret the curvilinear shape of this relationship as the effect of a balance between two mechanisms. Without overland dispersal, gradual differentiation according to an isolation by distance pattern occurred (results not shown). It results that nodes located close to the outlet and in distant branch tended to differentiate because they were located far apart. In contrast, genetic differentiation was strongly reduced when overland dispersal was allowed, i.e. when regular exchanges of migrants between branches could occur. However,  $F_{ST}$  increased again for high values of  $p$  when most migrants dispersed out-of-the network. Under such conditions, migration is expected to occur mostly within sub-units (sub-catchment) of the network. Accordingly, in-stream dispersal was very low, reducing the effect of effective migration along the network. Stronger population differentiation for low and high values of  $p$  is thus likely to be accounted for by the relative effects of the two dispersal modalities.

## *Perspectives and current limits of investigations of dispersal in river networks*

We developed an individual-based model to predict spatial genetic structure of organisms able to disperse in-stream and out-of-the network. We believe that this study can be useful to improve our understanding of population processes in river networks and we hope it will help stimulating other investigations on the topic. Nevertheless, we have to keep in mind that synthetic networks still represent simpler versions of natural river networks. For this reason, the effects on ecological processes of some of investigated factors, like network geometry, might be over- or under-estimated. Owing to computation constraints, our synthetic networks

represented only a part of the range of all existing networks. Much larger values of  $R_B$  can be observed for larger catchments. Moreover, we found that most variables interacted to affect the overall population differentiation. Thus, we cannot extrapolate our findings to higher levels of complexity yet. We focused on a restricted range of  $R_B$  values. Even if this range encompasses a large proportion of catchments observed, we are still unable to predict the shape of the surface response for very complex networks.

We used dichotomic networks and the bifurcation angle  $\alpha$  was fixed, i.e. all branches split with the same angle. Dichotomy is not an absolute rule in river network geometry and it would be highly interesting to consider the influence of this geometric factor in future studies. Moreover,  $\alpha$  is far from being constant. The variance of  $\alpha$  may be of some importance and ought to be considered too, even if we do not expect that drastically different conclusions would be drawn as the bifurcation angle was not the strongest effect we detected. Finally, natural river networks are subjected to stochastic factors (e.g. flood, drought, pollution) affecting intrinsic population growth rates, carrying capacities and deme extinction probability. These factors were not considered in our simulations. We kept the demographic part of the simulation quite simple. However, there are few reasons for demographic stochasticity such as local extinctions and re-colonisations to reduce overall population differentiation. Demographic stochasticity is usually shown to work the other way round as it reduces metapopulation effective size (Whitlock and Barton, 1997).

Considering overland dispersal opens new questions. For instance, river networks are modelled as objects completely isolated from external influences. In particular, no exchanges of migrants between neighboring catchments or other water bodies are allowed. In real situations, the closest branch to a river network can belong to a completely different catchment. If it makes no difference for strictly aquatic species, this situation has to be considered for species able of out-of-network movements. Such a situation is more likely to be observed in higher order segments (upstream parts) and is more likely in flat areas where terrestrial barriers to dispersal maybe less stringent than in areas with strong elevation variations. Overland movements from one valley to another could be simply impossible for many species when large altitude changes are required.

It is obvious from these considerations that the field of investigation is still vast. The present study is a first attempt to study gene flow patterns in river networks when different dispersal modalities are available. It completes other simulation studies or methods that have been developed recently to analyze population genetic structure (Neuenschwander 2006,

Labonne et al. 2008). The development of predictive methods specifically designed to address this issue is still highly required. Such methods are obviously important for basic science but they are also badly needed for conservation purpose given the general degradation of freshwater habitats and the reduction and fragmentation of populations they harbor. We hope this work, by considering the possibility of different dispersal modalities, will contribute to improve our understanding of population processes in river networks.

## **Acknowledgments**

We thank Guillaume Chapron, Eric D'Arlhac and Charline Retore for their helpful assistance in model programming. We gratefully acknowledge Jean-Michel Richer (LERIA, University of Angers) for his help to run the 4320 simulations and Mirabelle Gouat for her insights into statistics. This study was supported by the Région Pays de la Loire.

## REFERENCES

- Avise, J.C., 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.
- Beck, M.L. and Pruett-Jones, S., 2002. Fluctuating asymmetry, sexual selection, and survivorship in male dark-winged damselflies. *Ethology*, 108:779-791.
- Béliste, M., 2005. Measuring landscape connectivity: the challenge of behavioral landscape ecology. *Ecology*, 88:1988-1995.
- Bilton, D.T., Freeland, J.R. and Okamura, B., 2001. Dispersal in freshwater invertebrates. *Ann Rev Ecol Syst*, 32:159-181.
- Bowler, D.E. and Benton, T.G., 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biol Rev*, 80:205-225.
- Brumfield, R.T., Beerli, P., Nickerson, D.A. and Edwards, S.V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol Evol*, 18:249-256.
- Bullock, J.M., Kenward, R.E. and Hails, R., 2002. Dispersal ecology. Blackwell Science, Oxford.
- Burnham, K.P. and Anderson, D.R., 2002. Model selection and multimodel inference: A practical information-theoretic approach. 2nd ed. Springer-Verlag, New York.
- Campbell Grant, E.H., Lowe, W.H. and Fagan, W.F., 2007. Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecol Lett*, 10:165-175.
- Ceddia, M.G., Bartlett, M. and Perrings, C., 2007. Landscape gene flow, coexistence and threshold effect: The case of genetically modified herbicide tolerant oilseed rape (*Brassica napus*). *Ecol Model*, 205:169-180.
- Chaput-Bardy, A., Lemaire, C., Picard, D. and Secondi, J., 2008. In-stream and overland dispersal across a river network influences gene flow in a freshwater insect, *Calopteryx splendens*. *Mol Ecol* 17:3496-3505.
- Charles, S., Bravo de la Parra, R., Mallet, J.P., Persat, H. and Auger, P., 2000. Annual spawning migrations in modelling brown trout population dynamics inside an arborescent river network. *Ecol Model*, 133:15-31.

426 Clobert, J., Danchin, E., Dhondt, A.A. and Nichols, J.D., 2001. Dispersal, New York, 452  
427 p.

428 Cook, B., Hughes, J.M. and Bunn, S.E., 2002. Population structure of the freshwater prawn  
429 *Macrobrachium australiense* in drainages of western Queensland. Australia. Freshwater Biol,  
430 12:2098-2112.

431 Cote, D., Kehler, D.G., Bourne, C. and Wiersma, Y.F., 2009. A new measure of  
432 longitudinal connectivity for stream networks. Land Ecol, 24:101-113.

433 Cushman, S.A., McKelvey, K.S., Hayden, J. and Schwartz, M.K., 2006. Gene flow in  
434 complex landscapes: Testing multiple hypotheses with causal modeling. Am Nat, 168:486-  
435 499.

436 Fagan, W.F., 2002. Connectivity, fragmentation, and extinction risk in dendritic  
437 metapopulations. Ecology, 83:3243-3249.

438 Fraser, D.J., Lippe, C. and Bernatchez, L., 2004. Consequences of unequal population size,  
439 asymmetric gene flow and sex-biased dispersal on population structure in brook charr  
440 (*Salvelinus fontinalis*). Mol Ecol 13:67-80.

441 Ganio, L.M., Torgerser, C.E. and Gresswell, R.E., 2005. A geostatistical approach for  
442 describing pattern in stream networks. Front. Ecol. Environ., 3:138-144.

443 Gardner, B., Sullivan, P.J. and Lembo, A.J., 2003. Predicting stream temperatures:  
444 geostatistical model comparison using alternative distance metrics. Can J Fish Aqu Sci,  
445 60:344-351.

446 Gibbs, H.L., Gibbs, K.L., Siebenmann, M. and Collins, L., 1998. Genetic differentiation  
447 among populations of the rare mayfly *Siphonisca aerodromia* Needham. J. N. Am. Benthol.  
448 Soc., 17:461-474.

449 Grandjean, F., Souty-Grosset, C., Raimond, R. and Holdich, D.M., 1997. Geographical  
450 variation of mitochondrial DNA between populations of the white-clawed crayfish  
451 *Austropotamobius pallipes*. Freshwater Biol, 37:493-501.

452 Gustafson, E.J. and Gardner, R.H., 1996. The effect of landscape heterogeneity on the  
453 probability of patch colonization. Ecology, 77:94-107.



454 Hanski, I., 1999. Metapopulation ecology. Oxford University Press, New York, 313 p.

455 Hernandez-Martich, J.D., Novak, J.D., Smith, J.M. and Johns, P.E., 1995. Genetic  
456 structure of mosquitofish populations in the Altamaha and Ogeechee drainages of Georgia:  
457 reporting an undescribed form in the Ocmulgee river. *Biochem Syst Ecol*, 23:617-625.

458 Horton, R.E., 1945. Erosional development of streams and their drainage basins,  
459 hydrographical approach to quantitative morphology. *Geol Soc Am Bull*, 56:275-370.

460 Hurwood, D.A. and Hughes, J.M., 1998. Phylogeography of the freshwater fish, *Mogurnda*  
461 *adspersa*, in streams of north-eastern Queensland, Australia: evidence for altered drainage  
462 patterns. *Mol Ecol*, 7:1507-1517.

463 Hurwood, D.A. and Hughes, J.M., 2001. Nested clade analysis of the freshwater shrimp,  
464 *Caridina zebra* (Decapoda: Atyidae), from north-eastern Australia. *Mol Ecol*, 10:113-126.

465 Kovats, Z.E., Ciborowski, J.H. and Corkum, L.D., 1996. Inland dispersal of adult aquatic  
466 insects. *Freshwater Biol*, 36:265-276.

467 Labonne, J., Ravigné, V., Parisi, B. and Gaucherel, C., 2008. Linking dendritic network  
468 structures to population demogenetics: The downside of connectivity. *Oikos*, 117:1479-1490.

469 Little, L.S., Edwards, D. and Porter, D.E., 1997. Kriging in estuaries: as the crow flies, or  
470 as the fish swims? *J Exp Mar Biol Ecol*, 213:1-11.

471 Lowe, W.H., 2002. Landscape-scale spatial population dynamics in human-impacted  
472 stream systems. *Environ Manage*, 30:225-233.

473 Lowe, W.H., Likens, G.E. and Power, M.E., 2006. Linking scales in stream ecology.  
474 *BioScience*, 56:591-597.

475 Macneale, K.H., Peckarsky, B.L. and Likens, G.E., 2005. Stable isotopes identify dispersal  
476 patterns of stonefly populations living along stream corridors. *Freshwater Biol*, 50:1117-1130.

477 McCullagh, P. and Nelder, J.A., 1983. Generalized Linear Models. Chapman and Hall,  
478 London.

479 McGlashan, D.J. and Hughes, J.M., 2000. Reconciling patterns of genetic variation with  
480 stream structure, earth history and biology of the Australian freshwater fish *Craterocephalus*  
481 *stercusmuscarum* (Atherinidae). *Mol Ecol*, 9:1737-1751.

482       McGlashan, D.J., Hugues, J.M. and Bunn, S.E., 2001. Within-drainage population genetic  
483       structure of the freshwater fish *Pseudomugil signifer* (Pseudomugilidae) in northern Australia.  
484       Can. J. Fish. Aquat. Sci., 58:1842-1852.

485       Meffe, G.K. and Vrijenhoek, R.C., 1988. Conservation genetics in the management of  
486       desert fishes. Cons Biol, 2:157-169.

487       Miller, M.P., Blinn, D.W. and Keim, P., 2002. Correlation between observed dispersal  
488       capabilities and patterns of genetic differentiation in populations of four aquatic insect species  
489       from the Arizona White Mountains, USA. Freshwater Biol, 47:1660-1673.

490       Neuenschwanger, S., 2006. AQUASPLATCHE: a program to simulate genetic diversity in  
491       populations living in linear habitats. Mol Ecol Notes, 6:583-585.

492       Nuismer, S.L., Thompson, J.N. and Gomulkiewicz, R., 2003. Coevolution between hosts  
493       and parasites with partially overlapping geographic ranges. J Evol Biol, 16:1337-1345.

494       Peckarsky, B.L., Hughes, J.M., Mather, P.B., Hillyer, M. and Encalada, A.C., 2005. Are  
495       populations of mayflies living in adjacent fish and fishless streams genetically differentiated?  
496       Freshwater Biol, 50:42-51.

497       Petersen, I., Masters, Z., Hildrew, A.G. and Ormerod, S.J., 2004. Dispersal of adult aquatic  
498       insects in catchments of differing land use. J Appl Ecol, 41:934-950.

499       Power, M.E. and Dietrich, W.E., 2002. Food webs in river networks. Ecol Res, 17:451-  
500       471.

501       Rissler, L.J., Wilbur, H.M. and Taylor, D.R., 2004. The influence of ecology and genetics  
502       on behavioral variation in salamander populations across the Eastern Continental Divide. Am  
503       Nat, 164:201-213.

504       Ronce, O., 2007. How does it feel to be like a rolling stone? Ten questions about dispersal  
505       evolution. Ann Rev Ecol Syst, 38:231-253.

506       Rosso, R., Bacchi, B. and La Barbera, P., 1991. Fractal relation of mainstream length to  
507       catchment area in river networks. Water Resour Res, 27:381-387.

508       Sawyer, S., 1978. Isotropic random walks in a tree. Z. Wahrscheinlichkeitstheorie  
509       verwandle Gebiete, 42:279-292.

510 Schmidt, S.K., Hughes, J.M. and Bunn, S.E., 1995. Gene flow among conspecific  
 511 populations of *Baetis* spp. (Ephemeroptera): adult flight and larval drift. J. N. Am. Benthol.  
 512 Soc., 14:147-157.

513 Shaw, M.W., Harwood, T.D., Wilkinson, M.J. and Elliott, L., 2006. Assembling spatially  
 514 explicit landscape models of pollen and spore dispersal by wind for risk assessment. Proc R  
 515 Soc B Lond , 273:1705-1713.

516 Shimano. 2005. Characteritics of the stream networkcomposition of drainage basins in the  
 517 Japanese Islands. Environ.Geol. Water Sci. 20, 5-14.

518 Slatkin, M., 1985. Gene flow in natural populations. Ann Rev Ecol Syst, 16:393-430.

519 Strahler, A.N., 1957. Quantative analysis of watershed geomorphology. Transactions  
 520 AUG, 38:913-920.

521 Tarboton, D., Bras, R. and Rodriguez-Iturbe, I., 1990. Comment on "on the fractal  
 522 dimension of stream networks" by La Barbera and Rosso. Water Resour Res, 26:2243-2244.

523 Vuilleumier, S. and Fontanillas, P., 2007. Landscape structure affects dispersal in the  
 524 greater white-toothed shrew: Inference between genetic and simulated ecological distances.  
 525 Ecol Model, 201:369-376.

526 Ward, R.D., Woodward, M. and Skibinski, O.F., 1994. A comparison of genetic diversity  
 527 levels in marine, freshwater and anadromous fishes. J Fish Biol, 44.

528 Wang & Yin. 1997. A comparison of drainage networks derived from digital elevation  
 529 models at two scales, J. Hydrol. 210, 221-241.

530 Whitlock, M.C., Barton, N. H., 1997. The effective size of subdivided population. Genetics  
 531 146, 427-441.

532

## Figure captions

Figure 1: Bifurcation angle  $\alpha$  and dispersal pathways in river networks. In-stream dispersal is unaffected by  $\alpha$  (up). In contrast, overland dispersal is very sensitive to variation of  $\alpha$  as the fraction of the network receiving migrants from a node increases when  $\alpha$  decreases (down). Circles indicate the dispersal range and diamonds the nodes that can actually be reached from the focal node.

Figure 2: Examples of synthetic river network used for simulations in the present study. River order  $\omega$  and branch length ratio  $R_L$  were respectively 3 and 2. The bifurcation ratio  $R_B$  varied between 2 and 4. The figure shows only two values of the bifurcation angle  $\alpha$ , 15 and 60 degrees.

Figure 3: Variation of  $F_{ST}$  in river networks with different bifurcation angles  $\alpha$ , and levels of branching complexity  $R_B$ : (a)  $R_B=2$ ,  $\alpha=15^\circ$ , (b)  $R_B=2$ ,  $\alpha=90^\circ$ , (c)  $R_B=4$ ,  $\alpha=15^\circ$ , (d)  $R_B=4$ ,  $\alpha=90^\circ$ . Graphs show 20 replicates for each sets of simulation where  $\omega=3$  (river network order),  $R_L=2$  (the length ratio),  $m=0.1$  (dispersal rate),  $p=0.6$  (overland dispersal rate) are constant. Results indicate that migration-drift equilibrium is reached by the 300<sup>th</sup> generation. We used this value as the maximum number of iterations in the study.

Figure 4: : Cumulative probability of the number of nodes that can be reached from any site against the distance to the focal node. Functions are given for 3 levels of network complexity  $R_B$ , 2 bifurcation angles and two dispersal modalities (in-stream and overland). The vertical lines indicate the maximal dispersal distance in our simulations.

Figure 5: Response surface of overall  $F_{ST}$ -value to bifurcation angle  $\alpha$  and branching complexity  $R_B$ . Estimations were made using a weighted least square method on  $F_{ST}$  values for 20 replicates of each combination of  $\alpha$  and  $R_B$

562 Figure 6: Response surface of overall  $F_{ST}$  to bifurcation  $\alpha$ , and branching complexity  $R_B$ .  
563 Estimations were made using a weighted least square method on standard deviation of  $F_{ST}$   
564 values for 20 replicates of each combination of  $\alpha$  and  $R_B$ .

565

566 Figure 7: Interaction between  $p$ , the proportion of terrestrial dispersers, the bifurcation angle  
567  $\alpha$ , and network complexity  $R_B$ . on overall population differentiation  $F_{ST}$ . Triangles:  $\alpha=90^\circ$ ,  
568 diamonds  $\alpha=15^\circ$ .

569

570

**Table 1.** Generalized linear models after a stepwise AIC for the relationship between genetic differentiation ( $\log F_{ST}$ ), branching ratio of the river network ( $R_B$ ), bifurcation angle between branches ( $\alpha$ ), migration rate ( $m$ ), proportion of overland dispersal ( $p$ ), and the quadratic term ( $p^2$ ). For each model are shown the number of parameters ( $k$ ), log-likelihood (LL), change in Akaike's Information Criterion ( $\Delta AIC$ ), AIC weight ( $wAICc$ ), and the percentage deviance explained (%DE) in the  $\log F_{ST}$  response variable.

Model	Variables	k	Log-likelihood	$\Delta AIC$	AIC weights	%DE
1	complete model without the interaction $\alpha:m$	22	5305.831	0.0	0.514	86.51
2	complete model	23	5305.943	0.1	0.486	86.51

**Table 2.** ANOVA results of the most parsimonious GLM (see Model 1 in the Table 1). Df is the degree of freedom and Df Resid the degree of freedom of the residual deviance.

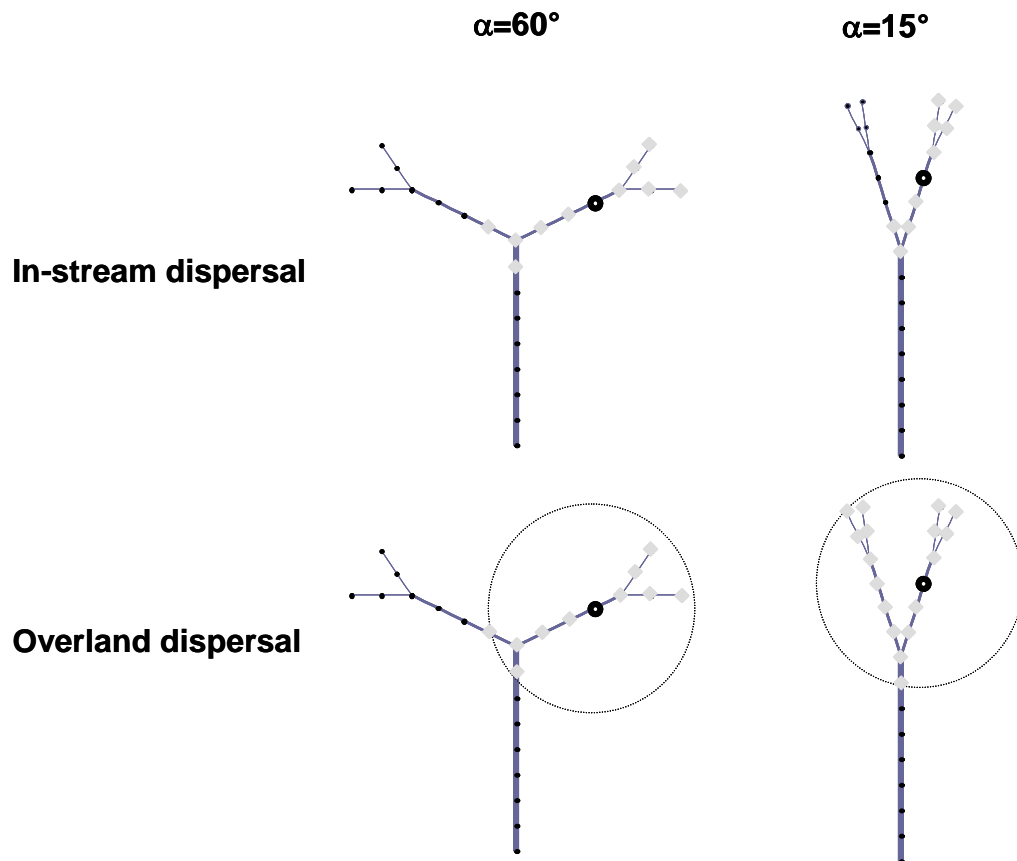
	Df	Residual deviance	Df Resid	Deviance	F	p
NULL			4319	160.801		
R <sub>B</sub>	2	3.999	4317	156.802	396.438	< 0.0001
$\alpha$	1	7.376	4316	149.427	1462.368	< 0.0001
m	1	<b>83.490</b>	4315	65.936	16553.809	< 0.0001
p	1	0.024	4314	65.912	4.795	0.02859
p <sup>2</sup>	1	<b>31.782</b>	4313	34.13	6301.440	< 0.0001
R <sub>B</sub> : $\alpha$	2	1.046	4311	33.084	103.697	< 0.0001
R <sub>B</sub> :m	2	0.094	4309	32.991	9.300	< 0.0001
R <sub>B</sub> :p	2	5.609	4307	27.382	556.011	< 0.0001
R <sub>B</sub> :p <sup>2</sup>	2	2.895	4305	24.487	286.962	< 0.0001
$\alpha$ :p	1	0.511	4304	23.976	101.387	< 0.0001
$\alpha$ :p <sup>2</sup>	1	0.795	4303	23.181	157.659	< 0.0001
m:p	1	0.030	4302	23.151	5.983	0.01448
m:p <sup>2</sup>	1	0.614	4301	22.537	121.695	< 0.0001
p:p <sup>2</sup>	1	0.850	4300	21.687	168.457	< 0.0001

**Table 3.** Effect of branching ratio of the river network ( $R_B$ ), bifurcation angle between branches ( $\alpha$ ), migration rate ( $m$ ), proportion of overland dispersal ( $p$ ), and the quadratic term ( $p^2$ ), on between genetic differentiation ( $\log F_{ST}$ ). Here is presented the most parsimonious Model 1 from the complete model. We used a Generalized Linear Model and applied a stepwise AIC with a backward selection procedure (see methods).  $R_B2$  and  $m10$  were taken as reference.

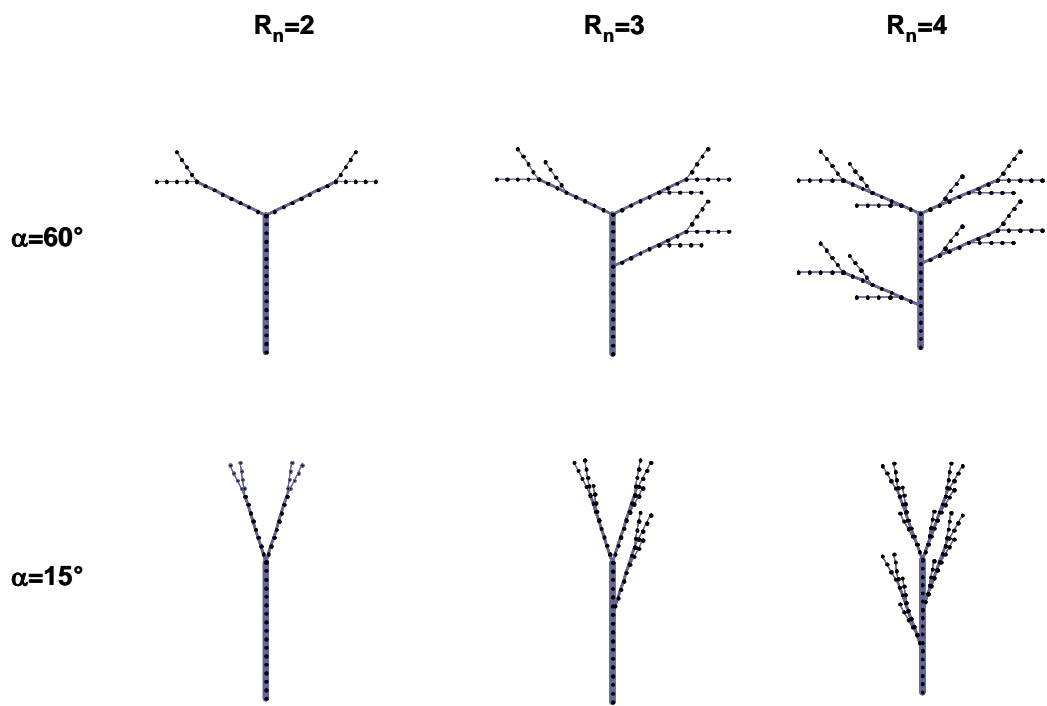
	Estimate	SE	t	p
(Intercept)	-9.63E-01	7.77E-03	-123.99	< 0.0001
$R_B3$	1.30E-01	8.42E-03	15.397	< 0.0001
$R_B4$	1.89E-01	8.42E-03	22.471	< 0.0001
$\alpha$	-3.71E-04	1.11E-04	-3.343	0.00084
$m20$	-2.52E-01	5.69E-03	-44.234	< 0.0001
$p$	-8.73E-03	3.94E-04	-22.166	< 0.0001
$p^2$	1.45E-04	6.90E-06	21.048	< 0.0001
$R_B3:\alpha$	5.00E-04	1.03E-04	4.843	< 0.0001
$R_B4:\alpha$	1.46E-03	1.03E-04	14.167	< 0.0001
$R_B3:m20$	2.03E-02	5.29E-03	3.829	0.00013
$R_B4:m20$	1.92E-02	5.29E-03	3.634	0.00028
$R_B3:p$	-6.86E-03	2.76E-04	-24.829	< 0.0001
$R_B4:p$	-8.22E-03	2.76E-04	-29.743	< 0.0001
$R_B3:p^2$	5.35E-05	2.65E-06	20.149	< 0.0001
$R_B4:p^2$	5.65E-05	2.65E-06	21.297	< 0.0001
$\alpha:p$	6.55E-05	4.40E-06	14.876	< 0.0001
$\alpha:p^2$	-5.31E-07	4.23E-08	-12.556	< 0.0001
$m20:p$	-2.54E-03	2.26E-04	-11.275	< 0.0001
$m20:p^2$	2.39E-05	2.17E-06	11.032	< 0.0001
$p:p^2$	-5.33E-07	4.11E-08	-12.979	< 0.0001



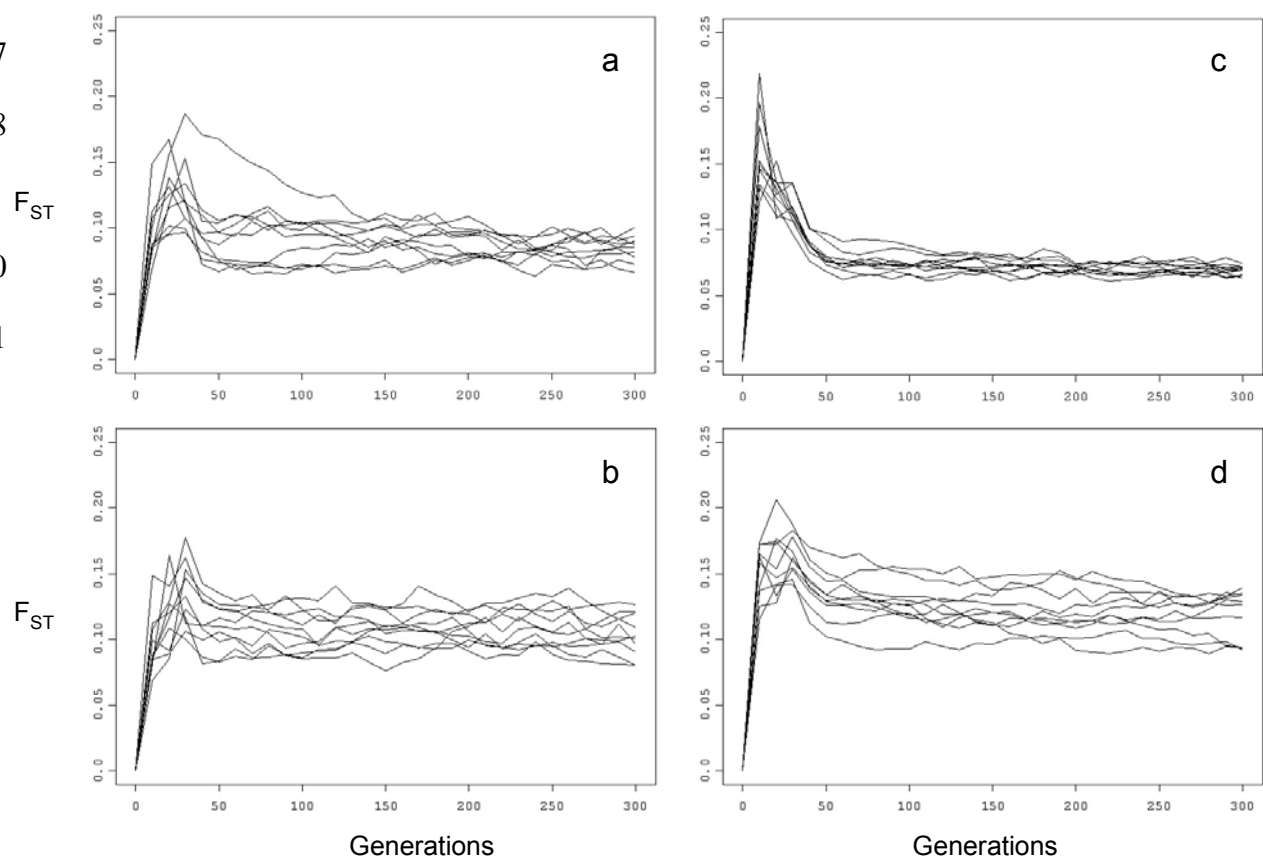
**Figure 1**



**Figure 2**



**Figure 3**

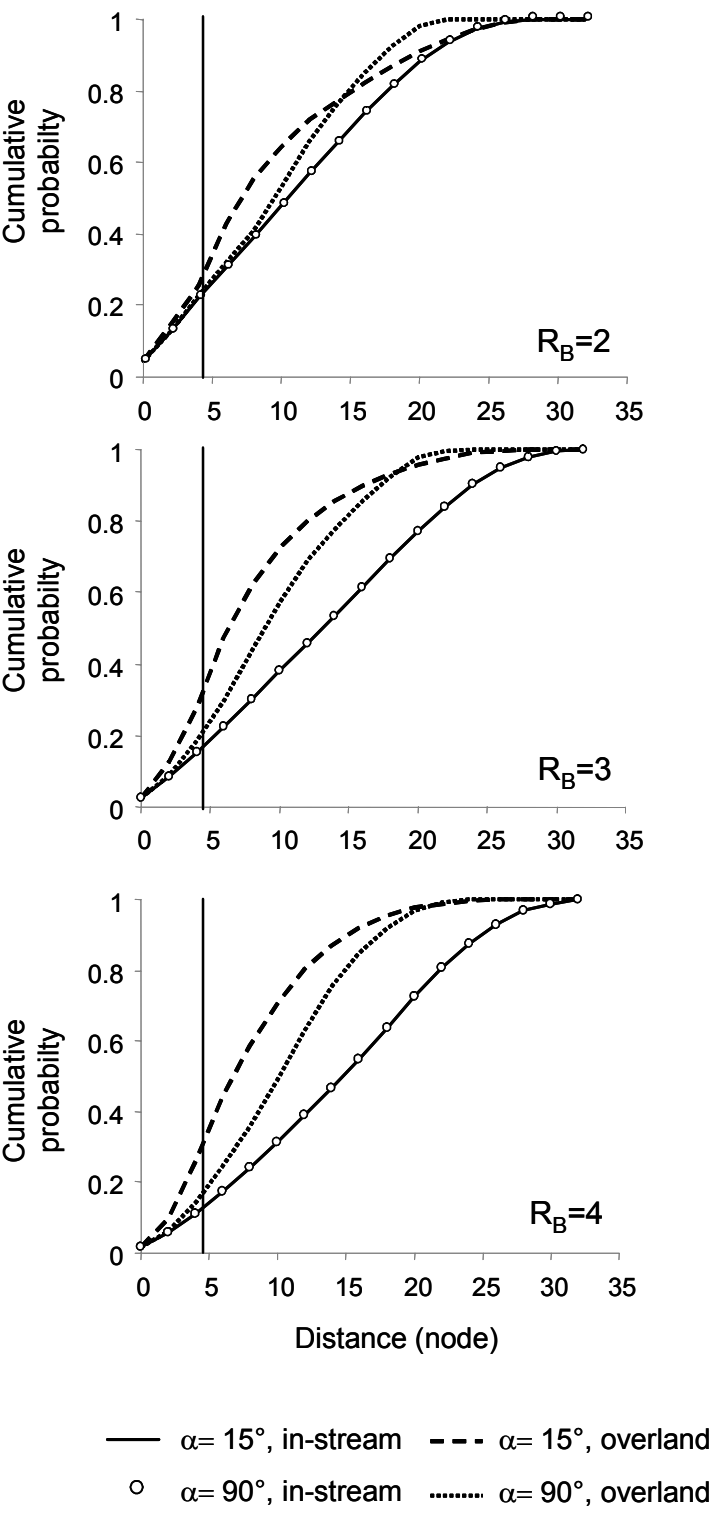


662 **Figure 4**

663

664

665

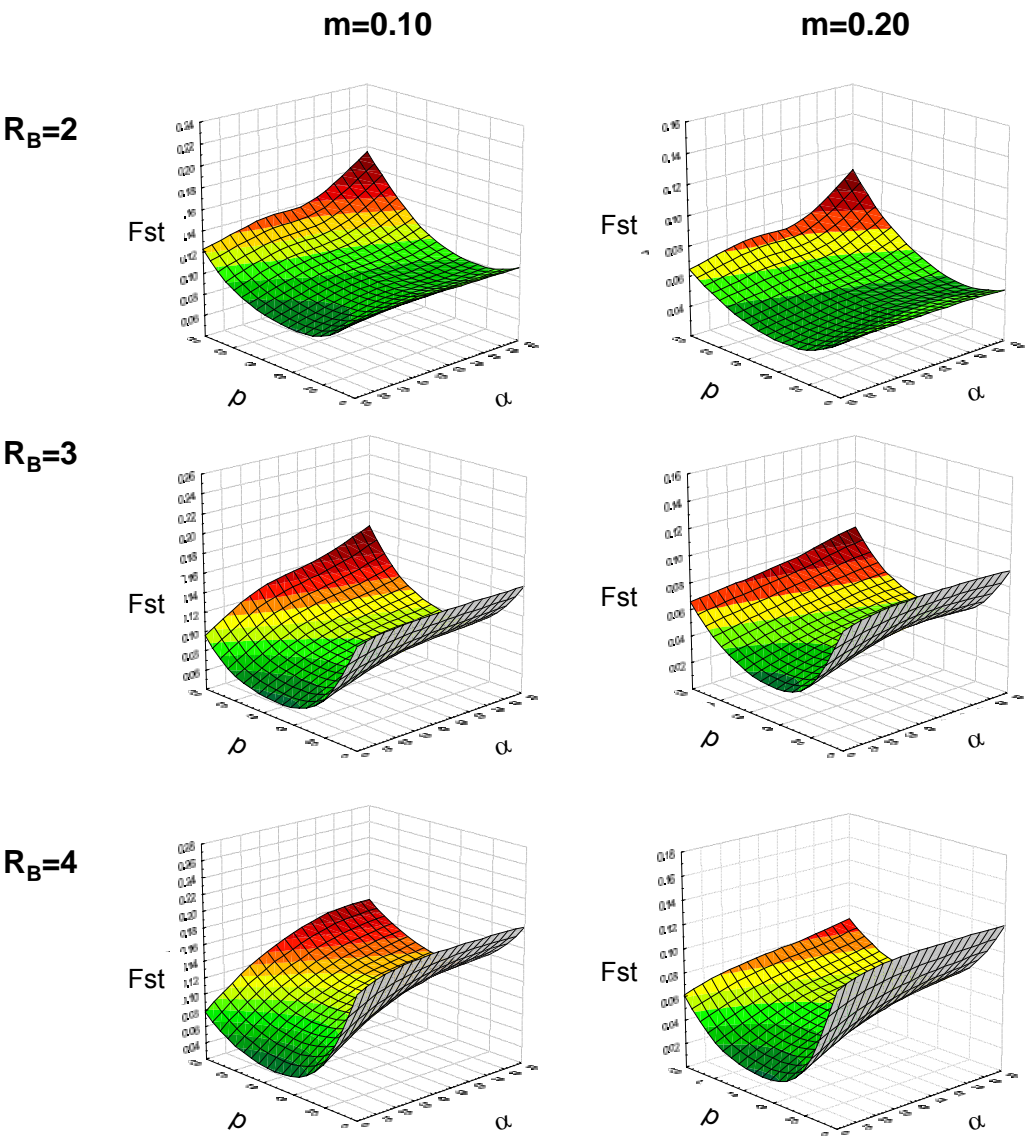


666 **Figure 5**

667

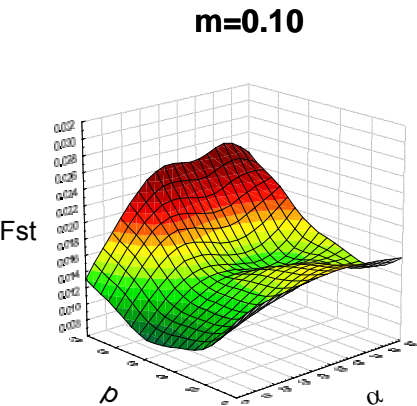
668

669

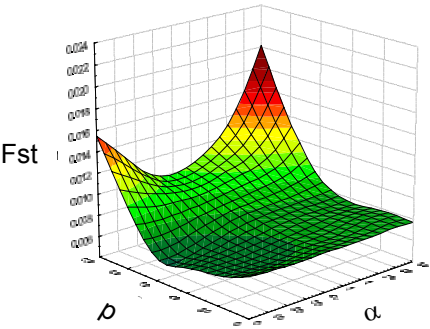


**Figure 6**

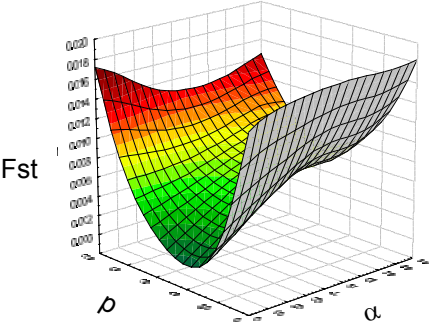
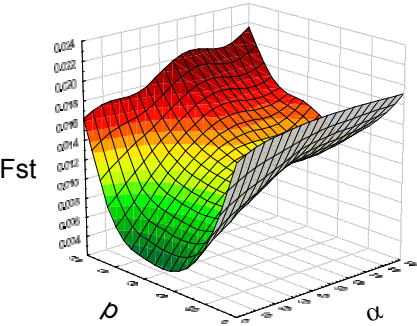
**$R_B=2$**



**$m=0.20$**



**$R_B=3$**



**$R_B=4$**

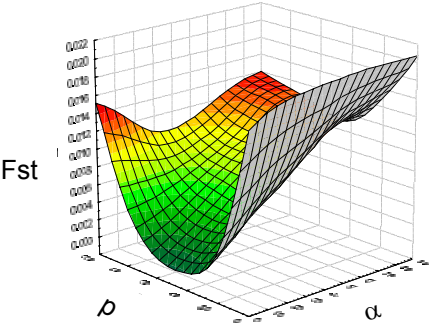
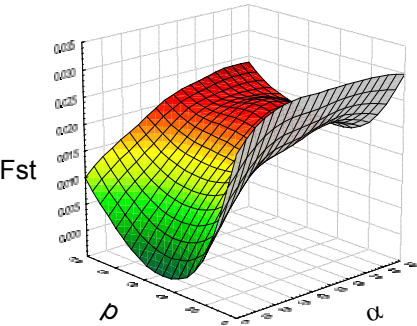


Figure 7

